

OM nucleic - nucleic search, using sw model
Run on: December 5, 2003, 07:48:44 ; Search time 2892 Seconds
(without alignments)
10439.594 Million cell updates/sec
Title: US-09-786-715-5
Perfect score: 738
Sequence: 1 gcaccaggaattotttagt.....aaaaaaaaaaaaaaaaaaaaa 738
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :		GenEmbl:*	
1:	gb_ba:*	1:	gb_ba:*
2:	gb_htg:*	2:	gb_htg:*
3:	gb_in:*	3:	gb_in:*
4:	gb_ov:*	4:	gb_ov:*
5:	gb_ov:*	5:	gb_ov:*
6:	gb_pat:*	6:	gb_pat:*
7:	gb_ph:*	7:	gb_ph:*
8:	gb_pl:*	8:	gb_pl:*
9:	gb_pr:*	9:	gb_pr:*
10:	gb_ro:*	10:	gb_ro:*
11:	gb_sts:*	11:	gb_sts:*
12:	gb_sy:*	12:	gb_sy:*
13:	gb_un:*	13:	gb_un:*
14:	gb_vl:*	14:	gb_vl:*
15:	em_ba:*	15:	em_ba:*
16:	em_fun:*	16:	em_fun:*
17:	em_hum:*	17:	em_hum:*
18:	em_in:*	18:	em_in:*
19:	em_mu:*	19:	em_mu:*
20:	em_om:*	20:	em_om:*
21:	em_or:*	21:	em_or:*
22:	em_ov:*	22:	em_ov:*
23:	em_pat:*	23:	em_pat:*
24:	em_ph:*	24:	em_ph:*
25:	em_pl:*	25:	em_pl:*
26:	em_ro:*	26:	em_ro:*
27:	em_sts:*	27:	em_sts:*

Result No.	Score	Match	Query Length	DB ID	Description
1	289.6	39.2	630	8	PSA310990
2	233.6	31.7	603	8	KCH10RNN
3	211.6	28.7	665	8	AV170650
4	203.4	27.6	698	8	NTRNA
5	201.8	27.3	480	8	AV088687
6	201.8	27.3	497	8	ANTHIOARA
7	198.6	26.9	345	6	AX505468
8	192	26.0	653	6	AR016869
9	192	26.0	653	6	AR020895
10	192	26.0	653	6	AR027218
11	192	26.0	653	6	AR038505
12	192	26.0	653	6	AR064647
13	192	26.0	653	6	AR067572
14	192	26.0	653	6	I38524
15	192	26.0	653	6	I56999
16	192	26.0	653	6	I59865
17	192	26.0	653	6	I75192
18	188	25.5	664	8	AF323593
19	186.6	25.3	784	8	AY271308
20	182.2	24.7	593	8	PSA319808
21	181	24.5	345	8	AF483265
22	173.2	23.5	580	8	BNUS380
23	172.8	23.4	590	8	AY088698
24	171.2	23.2	556	8	AK118542
25	170.6	23.1	360	6	AX505469
26	170.6	23.1	360	8	BT004710
27	170.4	23.1	528	8	ATTHRED4
28	170.4	23.1	530	8	AK118035
29	170.4	23.1	560	8	AY087159
30	170.4	23.1	561	8	AF360227
31	168.8	22.9	642	8	ATTHRED1
32	168.6	22.8	357	8	AY040028
33	158.8	21.5	740	8	AB053294

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

34 156.2 21.4 639 8 AV085117 Arabidops
35 156.6 21.2 357 6 AX412276
36 156.6 21.2 357 6 AX412277
37 156.6 21.2 357 6 AX412278
38 156.6 21.2 357 6 AX412279
39 156.6 21.2 357 6 AX412473
40 156.6 21.2 357 6 AX412474
41 156.6 21.2 357 6 AX412566
42 156.6 21.2 357 6 AX412917
43 156.6 21.2 357 6 AX507041
44 156.6 21.2 357 6 AX651357
45 156.6 21.2 402 8 AX114566 Arabidops

Search completed: December 5, 2003, 09:53:26
Job time : 2898 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 06:08:34 ; Search time 261 Seconds
(without alignments)
7089.627 Million cell updates/sec

Title: US-09-786-715-5

Perfect score: 738

Sequence: 1 gcaccaggaaatttttagt.....aaaaaaaaaaaaaaaa 738

Scoring table: IDENTITY NJC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq_19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	738	100.0	738	21	AAZ51739	Glycine max thione
2	208.8	28.3	402	25	ABX21664	Human GDP-mannose
3	201.8	27.3	561	21	AAC31522	Arabidopsis thalia
4	201.8	27.3	563	21	AAC34121	Arabidopsis thalia
5	201.6	27.3	3129	24	ABX53095	DNA encoding thior
6	201.6	27.3	3129	24	ABN9579	Phaseolin promoter
7	201.6	27.3	3888	24	ABX53096	DNA encoding Oleos
8	201.6	27.3	3888	24	ABN9580	Phaseolin promoter
9	199.2	27.0	4935	24	ABN9586	Promoter-Oleosin t
10	198.6	26.9	345	24	ABZ12358	Arabidopsis thalia
11	198.6	26.9	345	24	ABN9587	Arabidopsis thalia
12	197.4	26.7	3888	24	ABX53097	DNA encoding Thior
13	197.4	26.7	3888	24	ABN9581	Phaseolin promoter
14	195.6	26.5	470	21	AAC37781	Arabidopsis thalia
15	193.4	26.2	392	23	ABX18045	Human GDP-mannose
16	192	26.0	653	16	AAQ99753	Human SAR gene pl.
17	192	26.0	653	20	AAV62799	Tobacco SAR CHX in
18	192	26.0	653	20	AAV81683	Tobacco protein-sy
19	191.6	26.0	574	21	AAZ51738	Catalpa speciosa t
20	186.2	25.2	601	21	AAZ51740	Glycine max thione
21	176.6	23.9	390	25	ABX19403	Human GDP-mannose
22	175.6	23.8	328	21	AAH31097	Plant microsatelli
23	172.8	23.4	509	22	AAH87768	Peppermint plant o
24	172.8	23.4	590	21	AAC38792	Arabidopsis thalia
25	170.6	23.1	360	24	ABZ12359	Arabidopsis thalia
26	170.4	23.1	560	21	ABX41961	Arabidopsis thalia
27	170.2	23.1	614	21	AAZ51741	Vernonia mespilifo
28	169.8	23.0	419	21	AAH31942	Plant microsatelli
29	168.6	22.8	346	21	AAH31096	Plant microsatelli
30	167.6	22.7	320	21	AAH31785	Plant microsatelli
31	163	22.1	292	21	AAH31978	Plant microsatelli
32	158.8	21.5	356	21	AAH31804	Plant microsatelli
33	158.2	21.4	480	21	AAC36542	Arabidopsis thalia
34	157.2	21.3	353	21	AAH31118	Plant microsatelli
35	156.8	21.2	524	21	AAC33829	Arabidopsis thalia
36	156.6	21.2	357	24	ABZ13931	Arabidopsis thalia
37	156.6	21.2	652	21	AAC45656	Arabidopsis thalia
38	155.6	21.1	277	25	ABX26076	Human GDP-mannose
39	155.2	21.0	572	21	AAC52069	Arabidopsis thalia
40	155	21.0	563	21	AAC34211	Arabidopsis thalia
41	153.6	20.8	870	21	AAZ51737	Memordica charanti
42	152.8	20.7	328	21	AAH31222	Plant microsatelli
43	152.2	20.6	274	21	AAH31859	Plant microsatelli
44	151.8	20.6	576	25	ABX56868	Arabidopsis thalia
45	149.4	20.2	647	21	AAC38520	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
AAZ51739	
ID	AAZ51739 standard; cDNA; 738 BP.
XX	
AC	AAZ51739;
XX	
DT	04-JUL-2000 (first entry)
XX	
DE	Glycine max thioredoxin cDNA-1.
XX	
KW	Glycine max thioredoxin; clone sahlc-pk001.117; chimeric gene; soybean;
KW	transgenic plant; seed storage protein; allergenicity; ss.
XX	
OS	Glycine max.
XX	
FR	Key Location/Qualifiers
FT	CDS 68..430
FT	/*tag= a
FT	/product= "Thioredoxin"
XX	
PN	W0200014239-A2.
XX	
PD	16-MAR-2000.
XX	
PF	07-SEP-1999; 99WO-US20420.
XX	
PR	08-SEP-1998; 98US-0099501.
XX	
PA	(DUPO) DU FONT DE NEMOURS & CO E I.
XX	
PI	Allen SM, Thorpe CJ, Lu AL;
XX	
DR	WPI; 2000-236987/22.
DR	P-PSDB; MAY70481.
XX	
PT	New isolated polynucleotide encoding thioredoxin polypeptide is useful
PT	for producing transgenic plants with an altered level of thioredoxin: -
XX	
PS	Claim 3; Page 30; 33pp; English.
XX	
CC	The present cDNA sequence encodes Glycine max thioredoxin protein. The
CC	cDNA was derived from clone sahlc-pk001.117, which was isolated from a
CC	cDNA library prepared from soybean tissue sprayed with authority
CC	herbicide. Chimeric genes encoding all or a portion of the thioredoxin
CC	protein, in sense or antisense orientation are constructed, wherein
CC	expression of the chimeric gene results in production of altered levels
CC	of the thioredoxin protein in a transformed host cell. Thioredoxin is
CC	involved in the disassembly of seed storage proteins during germination of
CC	by reducing S-S bonds and in the bread making process. Over expression of
CC	thioredoxin in cereals may reduce the allergenicity of any transgenic
CC	protein engineered into cereal crops with high sulfhydryl content.
XX	
SQ	Sequence 738 BP; 260 A; 120 C; 146 G; 212 T; 0 other;

Query Match 100.0%; Score 738; DB 21; Length 738;
Best Local Similarity 100.0%; Pred. No. 3e-101;

Matches 739; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGAAATCTTTAGTTGTAACGTGACAAAGTTTCTGAGAAATAGAGATTAAT 60
DB 1 GCACGAGAAATCTTTAGTTGTAACGTGACAAAGTTTCTGAGAAATAGAGATTAAT 60

QY 61 TGAGAGATGGCTGGCTCATCGGAAGAGGACAGTCATTAGCTGCCACACCGTTGAAGA 120
DB 61 TGAGAGATGGCTGGCTCATCGGAAGAGGACAGTCATTAGCTGCCACACCGTTGAAGA 120

QY 121 ATGGAAGGATCACTCCAGAGGGCAACGATCCAGAACTCATTTGTGTGGATTATAC 180
DB 121 ATGGAAGGATCACTCCAGAGGGCAACGATCCAGAACTCATTTGTGTGGATTATAC 180

QY 181 TGCTTCTTGGGTGGACCATGCGTTTCTATGACCATCTCTTGGCTGAGCTGGCTAAGAA 240
DB 181 TGCTTCTTGGGTGGACCATGCGTTTCTATGACCATCTCTTGGCTGAGCTGGCTAAGAA 240

QY 241 GTTCACAGTGTCTATTCCTAAAGGTGGATGGAGCAATTAAGAGTGTCTTCACAGA 300
DB 241 GTTCACAGTGTCTATTCCTAAAGGTGGATGGAGCAATTAAGAGTGTCTTCACAGA 300

QY 301 TTGGGCTATGAGGCTATGCCACATTTTGTGTGTGTGTAAGAGGAGACGTTCTGGACAA 360
DB 301 TTGGGCTATGAGGCTATGCCACATTTTGTGTGTGTGTAAGAGGAGACGTTCTGGACAA 360

QY 361 AGTGGTGGGCAAGAGGATGAGCTGCAGCAGAAATACAGAAACATGTGGCTTCAGC 420
DB 361 AGTGGTGGGCAAGAGGATGAGCTGCAGCAGAAATACAGAAACATGTGGCTTCAGC 420

QY 421 TAGTGTCTTAATCTAGCTCAGCTTCAGAACTTTATATATGGCTTTCTTTTCAATCTT 480
DB 421 TAGTGTCTTAATCTAGCTCAGCTTCAGAACTTTATATATGGCTTTCTTTTCAATCTT 480

QY 481 GTACTAGACTTATGTTGGTATTTCTGTATTGCACCATCAGCTTTTCAAAGGTGATGAC 540
DB 481 GTACTAGACTTATGTTGGTATTTCTGTATTGCACCATCAGCTTTTCAAAGGTGATGAC 540

QY 541 TCCTATCATCTATTTCTGAATAGTAGTAAGTGGTCTTTCTTCGTTCTTAATAATAGTG 600
DB 541 TCCTATCATCTATTTCTGAATAGTAGTAAGTGGTCTTTCTTCGTTCTTAATAATAGTG 600

QY 601 GATGGTGTATATCATGATCTTAATATACATAGACCTTCCTGTTTCCCTTTAGTATTA 660
DB 601 GATGGTGTATATCATGATCTTAATATACATAGACCTTCCTGTTTCCCTTTAGTATTA 660

QY 661 AA 720
DB 661 AA 720

QY 721 AAAAAAAAAAAAAAAAAAAAAA 738
DB 721 AAAAAAAAAAAAAAAAAAAAAA 738

Search completed: December 5, 2003, 09:04:57
Job time : 284 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 09:53:30 ; Search time 318 Seconds
(without alignments)
7713.259 Million cell updates/sec

Title: US-09-786-715-5
Perfect score: 738
Sequence: 1 gcaaccaggaatttttagt.....aaaaaaaaaaaaaaaaaaaaa 739

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:**
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1 208.8 28.3 402 10 US-09-878-574-3723
2 201.6 27.3 3129 9 US-09-897-898-5
3 201.6 27.3 3129 12 US-10-032-201B-14
4 201.6 27.3 3888 9 US-09-897-898-7
5 201.6 27.3 3888 12 US-10-032-201B-16
6 198.2 27.0 4935 12 US-10-032-201B-33
7 198.6 26.9 345 10 US-09-938-842A-163
8 198.6 26.9 345 12 US-10-032-201B-38
9 197.4 26.7 3888 9 US-09-897-898-10
10 197.4 26.7 3888 12 US-10-032-201B-19
11 193.4 26.2 392 10 US-09-878-574-104
12 176.6 23.9 390 10 US-09-878-574-1462
13 170.6 23.1 360 10 US-09-938-842A-164
14 156.6 21.2 357 10 US-09-938-842A-1736
15 156.6 21.1 377 14 US-10-323-362-1
16 155.6 21.1 277 10 US-09-878-574-8135
17 144.4 19.6 370 10 US-09-878-574-3169
18 140.4 19.0 268 10 US-09-878-574-5995
19 140.2 19.0 540 12 US-10-349-782-12
20 139.2 18.9 382 12 US-10-091-841-3
21 139.2 18.9 393 12 US-10-091-841-5
22 138.2 18.7 260 10 US-09-878-574-12941
23 135.6 18.4 369 12 US-10-394-865-8
24 135.6 18.4 369 12 US-10-091-841-1
25 135 18.3 270 10 US-09-878-574-13445
26 130.8 17.7 272 10 US-09-878-574-12181
27 129.8 17.6 659 12 US-10-349-782-9
28 120 16.3 267 10 US-09-878-574-13137
29 105.2 14.3 402 10 US-09-938-842A-225
30 105.2 14.3 613 9 US-09-770-149-855
31 103.8 14.1 456 10 US-09-924-035A-677
32 100.6 13.6 297 12 US-10-349-782-11
33 94.6 12.8 250 9 US-09-923-876-1288
34 93.2 12.6 421 14 US-10-198-846-6007
35 92.2 12.5 476 12 US-09-814-353-4700
36 92.2 12.5 476 12 US-09-814-353-10999
37 88.4 12.0 470 11 US-09-918-995-22375
38 88.2 12.0 299 12 US-09-814-353-4844
39 88.2 12.0 299 12 US-09-814-353-11141
40 88 11.9 261 10 US-09-878-574-7974
41 87.2 11.8 419 10 US-09-960-352-11234
42 86.4 11.7 839 14 US-10-198-846-5829
43 86 11.7 2964 12 US-10-237-496-109
44 86 11.7 2964 12 US-10-242-074-109
45 86 11.7 2964 12 US-10-242-505-109

Search completed: December 5, 2003, 11:49:06
Job time : 321 sec

Sequence 3723, Ap
Sequence 5, Appli
Sequence 14, Appl
Sequence 7, Appli
Sequence 16, Appl
Sequence 33, Appl
Sequence 163, Appl
Sequence 36, Appl
Sequence 10, Appl
Sequence 19, Appl
Sequence 144, Appl
Sequence 1462, Ap
Sequence 164, App
Sequence 1736, Ap
Sequence 1, Appli
Sequence 8135, Ap
Sequence 3169, Ap
Sequence 5995, Ap
Sequence 12, Appl
Sequence 3, Appli
Sequence 5, Appli
Sequence 12941, A
Sequence 8, Appli
Sequence 13445, A
Sequence 12181, A
Sequence 9, Appli
Sequence 13137, A
Sequence 225, App
Sequence 855, App
Sequence 677, App
Sequence 11, Appl
Sequence 1288, Ap
Sequence 6007, Ap
Sequence 4700, Ap
Sequence 10999, A
Sequence 2375, A
Sequence 4844, Ap
Sequence 11141, A
Sequence 7974, Ap
Sequence 11234, A
Sequence 5829, Ap
Sequence 109, App
Sequence 109, App
Sequence 109, App

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 5, 2003, 10:29:10 ; Search time 1939 Seconds
(without alignments)
2531.797 Million cell updates/sec

Title: US-09-786-715-6
Perfect score: 615
Sequence: 1 MAGSSEGVVISCHVEMN.....GAKKDELQKQIKHVASASA 120

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2688711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-C=/cgr2.1/USPTO.spool.p/US09786715/runat_04122003_092646_3937/app.query.fasta_1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -ILOOPEXT=0
-UNITIS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09786715@cgrn.1.1.3745 @runat_04122003_092646_3937 -NCPU=6 -ICPU=3
-NO.MMAP -LARGEOQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LNGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: +
1: gb_ba: +
2: gb_hig: +
3: gb_in: +
4: gb_on: +
5: gb_ov: +
6: gb_pat: +
7: gb_ph: +
8: gb_pl: +
9: gb_pr: +
10: gb_ro: +
11: gb_sts: +
12: gb_sy: +

19	434	70.6	784	8	AY271308	Citrus x
20	429	69.8	357	8	AY040028	Arabidops
21	429	69.8	528	8	ATHHRE04	Z354762 A.thaliana
22	429	69.8	530	8	AK118035	Arabidops
23	429	69.8	530	8	AK118035	Arabidops
24	429	69.8	560	8	AY087159	Arabidops
25	429	69.8	561	8	AF360227	Arabidops
26	414	67.3	593	8	AF360227	Arabidops
27	388	63.1	526	11	GS3679	Populus t
28	388	63.1	601	8	OS092541	Oryza sativ
29	388	63.1	686	6	E08194	Oryza sativ
30	388	63.1	687	8	RICTH	RPS13 gene
31	385.5	62.7	590	8	AY088698	Arabidops
32	385	62.6	595	8	D87984	Fagopyrum e
33	383	62.3	366	6	AX654096	Sequence
34	381.5	62.0	360	6	AX050469	Sequence
35	381.5	62.0	360	6	BT004710	Arabidops
36	391.5	62.0	556	8	AK118542	Arabidops
37	381	62.0	745	8	AB053294	Oryza sat
38	378.5	61.5	642	8	ATHHRED1	Z354762 A.thaliana
39	376.5	61.2	357	6	AX412276	Sequence
40	376.5	61.2	357	6	AX412277	Sequence
41	376.5	61.2	357	6	AX412278	Sequence
42	376.5	61.2	357	6	AX412279	Sequence
43	376.5	61.2	357	6	AX412473	Sequence
44	376.5	61.2	357	6	AX412474	Sequence
45	376.5	61.2	357	6	AX412566	Sequence

ALIGNMENTS

RESULT 1	PSA310990	linear	630 bp	mRNA	PLN 12-APR-2001
LOCUS	PSA310990				
DEFINITION	Pisum sativum mRNA for thioredoxin h (trxn gene).				
ACCESSION	AJ310990				
VERSION	AJ310990.1	GI:13624883			
KEYWORDS	thioredoxin H; trxn gene.				
SOURCE	Pisum sativum (pea)				
ORGANISM	Pisum sativum				
REFERENCE	1				
AUTHORS	Traverso, J.A.				
JOURNAL	Thesis (2001) Department of Bioquímica, Biología Celular y Molecular de Plantas, Estación Experimental del Zaidin (C.S.I.C.), Granada, Spain				
REFERENCE	2				
AUTHORS	Traverso, J.A., Cazalis, R., Sahravy, M., Lopez-Gorge, J. and Chueca, A.				
TITLE	Sequencing, cloning and expression of Pisum sativum Thioredoxin h				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 630)				
AUTHORS	Traverso, J.A.				
TITLE	Direct Submission				
JOURNAL	Submitted (11-APR-2001) Traverso J.A., Bioquímica Biología Celular				

Result No.	Score	Query Match	Length	DB	ID	Description
13:	gb_uni.*					
14:	gb_vi.*					
15:	em_ba.*					
16:	em_fun.*					
17:	em_hum.*					
18:	em_in.*					
19:	em_mu.*					
20:	em_cm.*					
21:	em_or.*					
22:	em_ov.*					
23:	em_pat.*					
24:	em_ph.*					
25:	em_pl.*					
26:	em_rc.*					
27:	em_sts.*					
28:	em_uni.*					
29:	em_vi.*					
30:	em_htg_hum.*					
31:	em_htg_inv.*					
32:	em_htg_other.*					
33:	em_htg_mus.*					
34:	em_htg_pln.*					
35:	em_htg_rod.*					
36:	em_htg_man.*					
37:	em_htg_vrt.*					
38:	em_sy.*					
39:	em_htgo_hum.*					
40:	em_htgo_mus.*					
41:	em_htgo_other.*					

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	544	88.5	630	8	PSX31099C	AJ310990 Pisum sat
2	490	79.7	603	8	RCH109RXN	Z70677 R.communis
3	482	78.4	345	6	AX50546E	AX505468 Sequence
4	482	78.4	480	8	AY086887	AY086887 Arabidops
5	482	78.4	497	8	ATPH0A8A	Z14084 A.thaliana
6	460	74.6	664	8	AF323593	AF323593 Prunus pe
7	456	74.1	698	8	NTRNA	X59527 N.tabacum m
8	453	73.7	665	8	AY170650	AY170650 Pisum sat
9	445	72.4	653	6	AR016869	AR016869 Sequence
10	445	72.4	653	6	AR020895	AR020895 Sequence
11	445	72.4	653	6	AR027218	AR027218 Sequence
12	445	72.4	653	6	AR038505	AR038505 Sequence
13	445	72.4	653	6	AR064647	AR064647 Sequence
14	445	72.4	653	6	AR067572	AR067572 Sequence
15	445	72.4	653	6	I38524	I38524 Sequence 10
16	445	72.4	653	6	I56999	I56999 Sequence 10
17	445	72.4	653	6	I59865	I59865 Sequence 10
18	445	72.4	653	6	I75192	I75192 Sequence 10

y Molecular de plantas, Estacion Experimental del Zaidin (CSIC),
Profesor Albarada 1, Granada, 18008, SPAIN

FEATURES
source

..630
/organism="Pisum sativum"
/mol_type="mRNA"
/variety="Lincoln"
/db_xref="taxon:3888"
/tissue_type="Leaves"
1..630
/gene="trxh"
36..398
/gene="trxh"
/function="Unknown"
/codon_start=1
/evidence=experimental
/product="thioredoxin h"
/protein_id="CAC36986.1"
/db_xref="GI:1362484"
/translation="MAGSSEGVVISCHSDVNDILHRGNSKKLIVVDFASWGP
CRFIAPLGLAKFTNVIPLKVDVDELKSVQDWAVEAMTFVFKETILGWGA
KKEELQQTIERHVASNA"
398..609
/gene="trxh"
609
/gene="trxh"

3'UTR

polyA_site

BASE COUNT 195 a 95 c 141 g 199 t
ORIGIN

Alignment Scores:

Pred. No.: 4.38e-55 Length: 630
Score: 544.00 Matches: 102
Percent Similarity: 95.00% Conservative: 12
Best Local Similarity: 85.00% Mismatches: 6
Query Match: 88.46% Indels: 0
DB: 8 Gaps: 0

US-09-786-715-6 (1-120) x PSA310890 (1-630)

QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTtpAan 20
DB 36 ATGGCAGGTTCATCAGAGAGGACACAGTTATCAGCTGTACAGCGTGAATGATGGAAC 95
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40
DB 96 GATATCCTTCACAGAGGCAATGAATCCAGAACTGATTTGTGGACTTTACTGCTCT 155
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysPheThr 60
DB 156 TGGTGTGGACATGCGCTTTCATTCACACCAATCCTTGGTGAATGGCTAAGAGTTTACA 215
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTspAla 80
DB 216 AATGCTAATTCCTTAAGTGGACGTGGAGCACTAAGCTGCTGCTCAAGATGGGCT 275
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
:::|||||

Db 276 GTTAGGGCTATGCAACATTGGTTGTGTAAGAGAGAACGATTGGGCAAGTGTT 335
QY 101 GYAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
|||||:::|||||
Db 336 GGAGCAAGAAAGAGACTGCACAGACAAATTCAGAGACATGTGGCTTCACTATATGCT 395

RESULT 2

RCTHIORXN
LOCUS R.communis mRNA for thioredoxin. 603 bp mRNA linear PLN 04-APR-1996
DEFINITION R.communis mRNA for thioredoxin.
ACCESSION Z70677
VERSION Z70677.1 GI:1255953
KEYWORDS thioredoxin.
SOURCE Ricinus communis (castor bean)
ORGANISM Ricinus communis

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Ricinus.

REFERENCE 1 (bases 1 to 603)
AUTHORS Szederkenyi, J., Dolgener, B. and Schobert, C.
TITLE cDNA expressed in Ricinus cotyledons
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 603)
AUTHORS Dolgener, B.
TITLE Direct Submission
JOURNAL Submitted (04-APR-1996) Dolgener E., University of Bayreuth,
Department of Plant Physiology, Universitaetsstr.30, Bayreuth,
Germany, 95440

FEATURES

Location/Qualifiers
1..603
/organism="Ricinus communis"
/mol_type="mRNA"
/cultivar="Sanguineus"
/db_xref="taxon:3988"
/clone="pEDRH01.8"
/tissue_type="cotyledon"
/clone_lib="Lambda Excll"
/dev_stage="Seedling"
23..379
/codon_start=1
/product="thioredoxin"
/protein_id="CAA94534.1"
/db_xref="GI:1255954"

BASE COUNT 179 a 101 c 157 g 166 t
ORIGIN

Alignment Scores:

Pred. No.: 1.01e-48 Length: 603
Score: 490.00 Matches: 89
Percent Similarity: 91.38% Conservative: 17
Best Local Similarity: 76.72% Mismatches: 10
Query Match: 79.67% Indels: 0
DB: 8 Gaps: 0

CDS

DELOQTIAKEMATAST"
/translation="MAAEQGVIGCHTVEANNEQLQKNDTKGLLYVDFASWGPGR
FIAPFLAKKLPNVTFLKVDVDELKTVAEWAVESMPTFMLEKGIKMDKVVYGA

US-09-786-715-6 (1-120) x RCTHIOKXJ (1-603)					
QY	4	SerSerGluGluG.YqlnValIleSerCysHisThrValGluclutrrPAsnAspG.nLeu	23		
D6	26	GCAAGAGAAGGGACAGTGCAGTCGGTGCCACTGTGAGGCATGGATGACAAATTG	85		
QY	24	GlnLysGlyAsnGluSerLysLysLeullelvalValAaspPhethrAlaSerTrpCysGly	43		
D6	86	CAGAGGGAAATCATACCAGGAGCTGATCGTTGTGAITTTACTGCTCATGGTGTGA	145		
QY	44	ProCysArgPheIleAlaproPheLeuLacluLeuAlaLysLysPheThrSerValIle	63		
D6	146	CCAAGCGGTTCATGCTCCCTTCCTGGTGAGTGGCGCAAGAACTGCCAAATGTTACC	205		
QY	64	PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlalleGluAla	83		
D6	206	TTCCTGAAGTGGATGTGGATGAACCTGAAGACGTGTGCTCACGAGTGGGCTGTGGATCA	265		
QY	84	MetProThrPheValPheValLysGluGlyThrLeuLeuAmpLysValValGlyAlalys	103		
D6	266	ATGCCAACCTTATGTCTCTGAAGAGGGGAAGATATGGACAAGGTGTGTGGAGCAAG	325		
QY	104	LysAspGluLeuGlnLnLysIleGlnLysHisValAlaSerAlaSer	119		
D6	326	AAGAGGAACTGCAGCAACTATAGCGAACACATGGCTACTGCTCC	373		
RESULT 3					
AX505468		345 bp	DNA	linear	FAT 27-SEP-2002
LOCUS					
DEFINITION		Sequence 163 from Patent WO0216655.			
ACCESSION		AX505468			
VERSION		AX505468.1	GI:23386705		
KEYWORDS					
SOURCE					
ORGANISM					
Arabidopsis thaliana (thale cress)					
Arabidopsis thaliana					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsi.					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
Harper,J.F., Krepis,J., Wang,X. and Zhu,T. Stress-regulated genes of plants, transgenic plants containing same, and methods of use Patent: WO 0216655-A 163 28-FEB-2002; The Scripps Research Institute (US) ; Syngenta Participations AG (CH)					
FEATURES					
source					
1..345					
/organism="Arabidopsis thaliana"					
/mol_type="genomic DNA"					
/db_xref="taxon:3702"					
BASE COUNT	90 a	68 c	90 g	97 t	
ORIGIN					
Alignment Scores:					
Pred. No.:					345
Score:					88

AUTHORS	TITLE	JOURNAL	COMMENT
<p> Broyer,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K. Direct Submission Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA </p>	<p> This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessments of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the WS or LAer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly. </p>		

FEATURES	source	Location/Qualifiers
		1. .480
		/organism="Arabidopsis thaliana"
		/mol_type="mRNA"
		/db_xref="taxon:3702"
		/clone="g189"
CDS		13. .357
		/codon_start=1
		/product="thioredoxin h"
		/protein_id="AA067008.1"
		/db_xref="GI:21617958"
		translation="MASEEGQVIACHVETWNEQLKANESKTLVAVDTASWGQPCR FIATFFADIKLAKLNLVFLKYDTTELKSVASDMAIQMPTFNE LKEGKILDKVVGAKK DELQSIIRAKHLA"
BASE COUNT	124 a	93 c 111 g 152 t
ORIGIN		

```

Alignment Scores:
Pred. No.:      6.83e-48      Length:      480
Score:          482.00      Matches:      88
Percent Similarity: 89.38%      Conservative: 13
Best Local Similarity: 77.88%      Mismatches:  12
Query Match:      78.37%      Indels:       0
DB:              8      Gaps:         0

US-09-786-715-6 (1-120) * XY086687 (1-450)

QY      4 SerSerGluGluGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeu 23
      ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      16 GCTTCGGGAAGGACCAAGTCATCGCTGCCACACCGTTGACATGAACGACGACGCTT 75
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      24 CblLysGlyAsnGluSerLysLeuIleValValAspPheThrAlaSerTrpCysGly 43
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      76 CAGAGAGCTAATGAATCCAAACCTCTGTGTGTGTTGATTTCAGGGCTCTCTGGTGTGA 135
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle 63

```

[illegible]

	Location/Qualifiers	
FEATURES	1..457	/organism="Arabidopsis thaliana"
source		/mol_type="mRNA"
		/db_xref="taxon:3702"
CDS	30...374	
		/codon_start=1
		/product="Thioredoxin H"
		/protein_id="CAA78462.1"
		/db_xref="GI:116552"
		/db_xref="SWISS-PROT:P29448"
		/translation="MASSEGVAVCHITVETWNEQLQKANESKLTVVDFASWGQPCR FIAPFADLAKNKPNVLFLKVDTDELKSVASIDMAIQAMETMFLKEGKILDKVVAGAK DEIQSTIAKHLLA"
BASE COUNT	135 a	94 c 116 g 152 t
ORIGIN		

13: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1992.DAT:*
 14: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1993.DAT:*
 15: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1994.DAT:*
 16: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1995.DAT:*
 17: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1996.DAT:*
 18: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1997.DAT:*
 19: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1998.DAT:*
 20: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1999.DAT:*
 21: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2000.DAT:*
 22: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2001A.DAT:*
 23: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2001B.DAT:*
 24: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2002.DAT:*
 25: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	615	100.0	738	21	AAZ51739	Glycine max thior
2	485	78.9	3888	24	ABSS3097	INA encoding thior
3	485	78.9	3888	24	ABN89581	Phaseolin promoter
4	482	78.4	345	24	ABZ12358	Arabidopsis thalia
5	482	78.4	345	24	ABN89587	Arabidopsis thalia
6	482	78.4	561	21	AAC31522	Arabidopsis thalia
7	482	78.4	563	21	AAC34121	Arabidopsis thalia
8	482	78.4	3129	24	ABSS3095	INA encoding thior
9	482	78.4	3129	24	ABN89579	Phaseolin promoter
10	482	78.4	3888	24	ABSS3096	INA encoding Oleos
11	482	78.4	3888	24	ABN89580	Phaseolin promoter
12	478	77.7	470	21	AAC37781	Arabidopsis thalia
13	477	77.6	4935	24	ABN89586	Promoter-Oleosin t
14	472	76.7	509	22	AAH87768	Peppermint plant o
15	445	72.4	653	16	AAQ99783	Plant SAR gene pl.
16	445	72.4	653	20	AAV62799	Tobacco SAR CHX in
17	445	72.4	653	20	AAV81683	Tobacco protein-sy
18	438	71.2	402	25	ABX21664	Human GDP-mannose
19	430	69.9	574	21	AAZ51738	Catalpa speciosa t
20	429	69.8	560	21	AAC41961	Arabidopsis thalia
21	424	68.9	392	25	ABX18045	Human GDP-mannose
22	410	66.7	601	21	AAZ51740	Glycine max thior
23	408	66.3	328	21	AAZ51097	Plant microsatelli
24	403	65.5	346	21	AAZ51096	Plant microsatelli
25	398	64.7	320	21	AAZ51785	Plant microsatelli
26	392	63.7	614	21	AAZ51741	Vernonia mespilifo
27	391	63.6	390	25	ABX19403	Human GDP-mannose
28	388	63.1	686	15	AAQ78205	Gene coding for pr
29	388	63.1	687	21	AAC66375	Rice thiorredoxin h
30	385.5	62.7	590	21	AAC38792	Arabidopsis thalia
31	381.5	62.0	360	24	ABZ12359	Arabidopsis thalia
32	376.5	61.2	357	24	ABZ13931	Arabidopsis thalia
33	376.5	61.2	480	21	AAC36542	Arabidopsis thalia
34	376.5	61.2	524	21	AAC33829	Arabidopsis thalia

ALIGNMENTS

RESULT 1
 AAZ51739
 ID AAZ51739 standard; cDNA; 738 BP.

XX AC AAZ51739;
 XX DT 04-JUL-2000 (first entry)
 XX Glycine max thiorredoxin cDNA-1.
 XX Glycine max thiorredoxin; clone sahlc.pk001.117; chimeric gene; soybean;
 KW transgenic plant; seed storage protein; allergenicity; ss.
 XX Glycine max.
 XX Key Location/Qualifiers
 ET CDS 68..430
 FT /*tag= a
 FT /product= "thiorredoxin"
 XX WO20014239-A2.
 XX 16-MAR-2000.
 XX 07-SEP-1999; 99WO-US20420.
 XX 08-SEP-1998; 98US-0099501.
 XX (DUPO) DU FONT DE NEMOURS & CO E I.
 XX Allen SM, Thorpe CU, Lu AL;
 XX WPI: 2000-256987/22.
 XX P-PSDE; AAY70451.
 XX New isolated polynucleotide encoding thiorredoxin polypeptide is useful
 PT for producing transgenic plants with an altered level of thiorredoxin -
 XX Claim 3; Page 30; 33pp; English.
 XX The present cDNA sequence encodes Glycine max thiorredoxin protein. The
 CC cDNA was derived from clone sahlc.pk001.117, which was isolated from a

35 376.5 61.2 563 21 AAC34211 Arabidopsis thalia
 36 376.5 61.2 652 21 AAC48656 Arabidopsis thalia
 37 376 61.1 292 21 AAA31978 Plant microsatelli
 38 368 59.8 393 17 AAT10451 Hard wheat thior
 39 368 59.8 393 21 AAC62457 Wheat thiorredoxin
 40 367 59.7 382 21 AAC62456 Wheat thiorredoxin
 41 367 59.7 384 17 AAT10450 Soft wheat thior
 42 367 59.7 419 21 AAA31942 Plant microsatelli
 43 367 59.7 572 21 AAC52069 Arabidopsis thalia
 44 366 59.5 870 21 AAZ51737 Monardica charanti
 45 364.5 59.3 576 25 ABX56868 Arabidopsis thalia

CC cDNA library prepared from soybean tissue sprayed with authority
CC herbicide. Chimeric genes encoding all or a portion of the thioredoxin
CC protein, in sense or antisense orientation are constructed, wherein
CC expression of the chimeric gene results in production of altered levels
CC of the thioredoxin protein in a transformed host cell. Thioredoxin is
CC involved in the disassembly of seed storage proteins during germination
CC by reducing S-S bonds and in the bread making process. Over expression of
CC thioredoxin in cereals may reduce the allergenicity of any transgenic
CC protein engineered into cereal crops with high sulphydryl content.
XX
SQ Sequence 738 BP; 260 A; 120 C; 146 G; 212 T; 0 other;

Alignment Scores:
Pred. No.: 4.7e-73 Length: 738
Score: 615.00 Matches: 120
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0
US-09-786-715-6 (1-120) x AA251739 (1-738)
QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20
DB 68 ATGGCTGGCTCATCGAAGAGGACAGTCATTCAGTCCACACCGTGAAGATGGAC 127
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40
DB 128 GATCAACTCCAGAGGCGACGATCCAGAACTCATTTGTGGATTTTACTGCTCT 187
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
DB 188 TGGTGTGGACCATCGCTTCATTCGACCATTCCTGGCTGAGCTGGCTAAGAGTTCACA 247
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
DB 248 AGTGTATATTCCTAAGGTGGATGTGACGAATTAAGAGTGTTCAGAGTTGGGCT 307
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
DB 308 ATTGAGGCTATGCCCACTTTTGTGTTTGTGAAGAGGGAACGCTTCGACAAAGTGTG 367
QY 101 GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
DB 368 GGAGCAAGAGGATGAGCTGCAGCAGAAATACAGAAACATGTGGCTTCAGTACTGCT 427

Search completed: December 5, 2003, 11:53:37
Job time : 258 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - nucleic search, using frame_plus_p2n model

Run on: December 5, 2003, 12:26:17 ; Search time 259 Seconds
(without alignments)
1539.892 Million cell updates/sec

Title: US-09-786-715-6
Perfect score: 615
Sequence: 1 MAGSSEGGVISHCTVEEWN.....GAKKQELQKTKQKVASASA 120
Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US09786715/runat_04122003.092649.4034/app.query.fasta_1.263
-DB=Published Applications NA -QPM=fastap -SUFFIX=p2n.cp30 -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSZF=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09786715 @CGN 1.1.85 @runat_04122003.092649.4034
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PTC_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubpna/PTCUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:

Sequence 3641, Ap
Sequence 1059, Ap
Sequence 1145, Ap

43 222 36.1 418 14 US-10-040-862-3641
c 44 222 36.1 464 10 US-09-998-598-1059
45 222 36.1 479 11 US-09-535-459-1145

Search completed: December 5, 2003, 14:00:05
Job time : 266 secs

13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:**
14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:**
15: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:**
16: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:**
17: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:**

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	485	78.9	3888	9	US-09-897-898-10	Sequence 10, Appl
2	485	78.9	3888	12	US-10-032-201B-19	Sequence 19, Appl
3	482	78.4	345	10	US-09-938-842A-163	Sequence 163, App
4	482	78.4	345	12	US-10-032-201B-38	Sequence 38, Appl
5	482	78.4	3129	9	US-09-897-898-5	Sequence 5, Appl
6	482	78.4	3129	12	US-10-032-201B-14	Sequence 14, Appl
7	482	78.4	3888	9	US-09-897-898-7	Sequence 7, Appl
8	482	78.4	3888	12	US-10-032-201B-16	Sequence 16, Appl
9	477	77.6	4935	12	US-10-032-201B-33	Sequence 33, Appl
10	438	71.2	402	10	US-09-878-574-3723	Sequence 3723, Ap
11	424	68.9	392	10	US-09-878-574-104	Sequence 104, App
12	391	63.6	390	10	US-09-878-574-1462	Sequence 1462, Ap
13	397	62.9	540	12	US-10-349-782-12	Sequence 12, Appl
14	381.5	62.0	360	10	US-09-938-842A-164	Sequence 164, App
15	376.5	61.2	357	10	US-09-938-842A-1736	Sequence 1736, Ap
16	376.5	61.2	357	14	US-10-323-362-1	Sequence 1, Appl
17	368	59.8	393	12	US-10-091-841-5	Sequence 5, Appl
18	367	59.7	382	12	US-10-091-841-3	Sequence 3, Appl
19	361	58.7	369	12	US-10-194-865-8	Sequence 8, Appl
20	361	58.7	369	12	US-10-091-841-1	Sequence 1, Appl
21	358	58.2	659	12	US-10-349-782-9	Sequence 9, Appl
22	316	51.4	277	10	US-09-878-574-8135	Sequence 8135, Ap
23	300	48.8	260	10	US-09-878-574-12941	Sequence 12941, A
24	293	47.6	370	10	US-09-878-574-3169	Sequence 3169, Ap
25	289	47.0	268	10	US-09-878-574-5995	Sequence 5995, Ap
26	285.5	46.4	402	10	US-09-938-842A-225	Sequence 225, App
27	285.5	46.4	613	9	US-09-770-149-855	Sequence 855, App
28	283	46.0	270	10	US-09-878-574-13445	Sequence 13445, A
29	279	45.4	272	10	US-09-878-574-12181	Sequence 12181, A
30	274	44.6	297	12	US-10-349-782-11	Sequence 11, Appl
31	273	44.4	2453	12	US-10-431-304-1	Sequence 1, Appl
32	272	44.2	456	10	US-09-924-035A-677	Sequence 677, App
33	256	41.6	250	9	US-09-923-876-1288	Sequence 1288, Ap
34	255	41.5	390	10	US-09-938-842A-2688	Sequence 2688, Ap
35	252	41.0	267	10	US-09-878-574-13137	Sequence 13137, A
36	237	38.5	261	10	US-09-878-574-7974	Sequence 7974, Ap
37	229	37.2	440	9	US-09-770-444-914	Sequence 914, App
38	226	36.7	357	10	US-09-938-842A-147	Sequence 147, App
39	222	36.1	301	9	US-09-615-349-76	Sequence 76, Appl
40	222	36.1	318	12	US-10-032-201B-46	Sequence 46, Appl
41	222	36.1	413	14	US-10-102-524-1396	Sequence 1396, Ap
42	222	36.1	418	10	US-09-796-692-3641	Sequence 3641, Ap

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 5, 2003, 11:18:30 ; Search time 2035 Seconds
(without alignments)
1419.239 Million cell updates/sec

Title: US-09-786-715-6
Perfect score: 615
Sequence: 1 MAGSSEGGVTSCHTVEEN.....CAKKDELOQKIQHVASA 120

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 100%
Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=cg2_1/USPTO_spool_p/US09786715/runat_04122003_092647_3948/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09786715 &CGN 1 1 2874 @runat_04122003_092647_3948 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=1.0 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*

- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fur:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rtd:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	615	100.0	459	12	BI699372 sag36c07.
2	615	100.0	495	13	BQ627882 sac65b02.
3	615	100.0	504	10	BG508580 sag74r06.
4	615	100.0	572	13	BQ630469 sag07f11.
5	615	100.0	562	13	BQ630383 sag06f10.
6	611	99.3	445	13	BQ578741 sar60h12.
c	7	600	97.6	14	CA937586 sar21g03.
	8	583	94.8	12	BE803058 sr3ic05.y
9	582	94.6	522	13	BQ786610 sag70f04.
	10	581	94.5	452	9 AY988470 ad02f07.y
11	581	94.5	466	9	AW569018 s174e02.y
	12	577	93.8	613	12 BG838879 Gc01_06f0
c	13	566	92.0	627	14 CD401101 Gm.ct2312
	14	556	90.4	687	14 CD038084 UTFFI002.
15	533	86.7	570	14	CA901886 PCSCL1470
	16	528	85.9	572	14 CA411141 E434.F -P
17	515	83.7	632	14	CA896875 PCF05423
	18	512	83.3	445	10 BE610138 sq62b12.y
19	512	83.3	512	14	CA794815 Cae.BJ.18
	20	502	81.6	498	14 CA901885 PCSG11713
21	501	81.5	311	12	BI094898 sag61e08.
	22	500	81.3	537	13 BQ625213 USDA-FP.0
23	500	81.3	655	14	CB291687 UCRCS01.0
	24	497	80.8	653	14 CB293759 UCRCS01.0
25	495	80.5	517	14	CA795737 Cae.BJ.27
	26	494	80.3	654	14 CB291686 UCRCS01.0
c	27	490	79.7	394	14 CA901887 PCSCL5426
	28	490	79.7	493	14 CB303530 wsa0001.1
29	490	79.7	546	10	BE053835 GA_Ea000
	30	485	78.9	480	13 BUB22062 UB328PB12

FEATURES
 source
 Location/Qualifiers
 1. 459
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl091-878"
 /tissue_type="Roots of 7 day old 'Bragg' seedlings"
 /dev_stages="7 days old"
 /lab_host="DH1CB"
 /clone_lib="Gm-cl08:"
 /notes="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; The mRNA was isolated from roots of 7 day old 'Bragg' seedlings that were mock-infected 48 hours prior to harvest. Dr. Gary Stacey generously donated the tissue. The roots were flash-frozen in liquid nitrogen. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGAGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (40U/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using Sephacryl S-500 High Resolution (Pharmacia Biotech) in a 2-mm diameter column and a bed volume of approximately 1ml. The column eluent was precipitated, redissolved, and ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). This library was constructed in the laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at Northern Arizona University."

BASE COUNT 132 a 81 c 114 g 132 t
 ORIGIN

Alignment Scores:
 Pred. No.: 6,61e-72 Length: 459
 Score: 615.00 Matches: 120
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-09-786-715-6 (1-120) x BI699372 (1-459)

QY 1 MetaGlySerSerGluGluGlnValIleSerOysHisThrValGluGluTrpAsn 20
 DB 33 ATGGCTGGCTCATCGAGAGAGGACCAAGTCATTAGCTGCCACACCGTGAAGTGAAC 92

485 78.9 555 13 BU875659 V009301 P
 CA825633 R63G02 tw
 BU863412 S027G05 P
 BU855457 M1480 pep
 CA825259 N58U03 tw
 CA824794 R48E09 tw
 CA824034 R35D03 tw
 CB077020 hj47g06.g
 AU300651 AU300651
 AW255195 M1185 pep
 BI892660 saf36c06.
 T24347 crs1466 lm
 CB377001 Hb02G03 L
 BE034543 GA_Ea002
 BM111010 EST558546

31 485 78.9 555 13 BU875659
 32 485 78.9 564 14 CA825633
 33 485 78.9 587 13 BU863412
 34 485 78.9 622 9 AW255457
 35 483 78.5 579 14 CA825259
 36 481 78.2 446 14 CA824794
 37 480 78.0 581 14 CA824034
 38 475.5 77.3 543 14 CB077020
 39 470 76.4 634 9 AU300651
 40 469 76.3 586 9 AW255195
 41 468 76.1 316 12 BI892660
 42 468 76.1 413 14 T24347
 43 467 75.9 540 14 CB377001
 44 467 75.9 602 10 BE034543
 45 466 75.8 523 12 BM111010

ALIGNMENTS
 459 bp mRNA linear EST 29-NOV-2001
 BI699372 Gm-cl091 Glycine max cDNA clone GENOME SYSTEMS CLONE
 seq36c07.yl Gm-cl091-878 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; ,
 ID: Gm-cl091-878 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; ,
 mRNA sequence.

BI699372
 seq36c07.yl Gm-cl091 Glycine max cDNA clone GENOME SYSTEMS CLONE
 ID: Gm-cl091-878 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; ,
 mRNA sequence.

BI699372 GI:15662001
 BI699372
 EST.
 Glycine max (soybean)
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 459)
 Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
 ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
 Wylie,T., Underwood,K., Streptoe,M., Theising,B., Allen,M., Bowers
 ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
 ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
 ,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 Unpublished
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Putative full length read
 vector to vector length is 527 This clone is available through:
 ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL
 35801 For further information call: (800)-533-4363 or contact via
 email: cdu@resgen.com
 Seq primer: -403R from Gibco
 High quality sequence stop: 424.

RESULT 1
 BI699372
 LOCUS
 DEFINITION
 accession
 version
 keywords
 source
 organism
 reference
 authors
 title
 journal
 comment

vector to vector length is 605
Seq primer: -40RP from Gibco
High quality sequence stop: 420.

FEATURES
source
1..495
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl073-3915"
/tissue_type="seedlings induced for symptoms of SDS
(Sudden Death Syndrome) disease"
/dev_stage="2-3 weeks old"
/lab_host="DH10B"
/clone_lib="Gm-cl073"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from 2-3 week old seedlings that were induced for symptoms
of SDS (Sudden Death Syndrome) disease by the
translocation of culture filtrate of *Fusarium solani* f.
sp. *glycines* (Plant Cell Report 18:375-380). Cultivar
Williams 82 is susceptible to the disease SDS. Plant
tissue (expanded leaves, folded leaves, and new shoots)
were collected at 1, 6, 24, and 48 hrs. after inoculation
and their mRNA pooled equally for cDNA construction. The
library was prepared using the Stratagene pBluescript II
SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with an XhoI restriction site. EcoRI
adaptors were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA insert is protected
from XhoI digestion via methylation during first strand
synthesis. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
E. coli ElectroMax DH10B host cells. Plants were inoculated
by Shuxian Li (Glen Hartman lab, University of Illinois).
Library was constructed by Reena Philip and Steve Clough
(Lila Vodkin lab, University of Illinois)."

BASE COUNT 138 a 88 c 120 g 149 t
ORIGIN

Alignment Scores:
Pred. No.: 7.31e-72 Length: 495
Score: 615.00 Matches: 120
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x BQ627882 (1-495)

Qy 1 MetAlaGlySerSerGluGlyGlnValIleSerCysHisThrValGluTrpAsn 20
|||||
Db 35 ATGGCTGGCTCAICGGAGAGGGACAGTCATTAGCTGCCACCGCTTGAAGATGGAC 94
|||||
Qy 21 AspGlnLeuGlnLysGlyAsnGlnSerLysLysLeuIleValValAspPheThrAlaSer 40
|||||

21 AspGlnLeuGlnLysGlyAsnGlnSerLysLysLeuIleValValAspPheThrAlaSer 40
|||||
Db 93 GATCACTCCAGAGGGCAACGATCCAGAACTCACTGTTGTGGATTTTACTGCTCT 152
|||||
Qy 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
|||||
Db 153 TGGTGTGGACCATCGCTTTCATTGCAACCTTCTTGGCTGAGCTGGCTAAGAGTCCAC 212
|||||
Qy 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
|||||
Db 213 AGTGTATATCTCTAAAGGTGGAGTGGAGCAATTAAGAGGTGTTCTCAGATGGGCT 272
|||||
Qy 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
|||||
Db 273 ATTGAGGCTATGCCACTTTTGTGTTGTGAAGAGGGAGCGCTCTCGACAAAGTGGTG 332
|||||
Qy 101 GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
|||||
Db 333 GGAGCAAGAAGGATGAGCTGCAGCGAGAAATACAGAAATCGCTTACGCTAGTGTCT 392
|||||

RESULT 2
BQ627882
LOCUS
DEFINITION
Gm-cl073-3915 5' similar to FR:Q43636 Q43636 THIOREDOXIN. ; mRNA
sequence.
ACCESSION
VERSION BQ627882
KEYWORDS
SOURCE EST.
ORGANISM
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
1 (bases 1 to 495)
Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: cou@resgen.com web site:
www.resgen.com
Putative full length read


```

/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl072-1043"
/tissue_type="seedlings induced for symptoms of SDS
(Sudden Death Syndrome) disease"
/dev_stages="2-3 weeks old"
/lab_host="DH10B"
/clone_lib="Gm-cl072"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from 2-3 week old seedlings that were induced for symptoms
of SDS (Sudden Death Syndrome) disease by the
translocation of culture filtrate of Fusarium solani f.
sp. glycines (Plant Cell Report 18:375-380). Cultivar PI
567374 is partially resistant to the disease SDS. Plant
tissue (expanded leaves, folded leaves, and new shoots)
were collected at 1, 6, 24, and 48 hrs. after inoculation
and their mRNA pooled equally for cDNA construction. The
library was prepared using the Stratagene pBluescript II
SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with an XhoI restriction site. EcoRI
adaptors were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA insert is protected
from XhoI digestion via methylation during first strand
synthesis. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
E.coli ElectroMax DH10B host cells. Plants were inoculated
by Shuxian Li (Glen Hartman lab, University of Illinois).
Library was constructed by Steve Clough (Lila Vocklin lab,
University of Illinois)."
BASE COUNT      143 a      123 g      150 t
ORIGIN
Alignment Scores:
Pred. No.:      7,49e-72      Length:      504
Score:          615.00      Matches:      120
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              10      Gaps:      0

US-09-766-715-6 (1-120) x B3508580 (1-504)

QY      1 MetAlaglySerSerGluGluGlyGlnValIleSerCysHisThrValGluItrpAsn 20
      |||||
Db      37 ATGGCTGGCTCATCGGAGGAGGAGCAAGTCATTAGTCGCCACACCGTTGAAGAATGAAC 96

QY      21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40
      |||||
Db      97 GATCAACTCCAGAGGAGGCAACGAAATCCAGAACTCATTTGTTGGATTTTACTGCTTCT 156

QY      41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
      |||||
Db      157 TGTGTGGACCATGGCGTTTCATTGCACCATCTTCGTGAGCTGGCTAGAGATTACA 216

QY      61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80

```

```

95 GATCAACTCCAGAGGAGGCAACGAAATCCAGAACTCATTTGTTGGATTTTACTGCTTCT 154
QY      41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
      |||||
Db      155 TGTGTGGACCATGGCGTTTCATTGCACCATCTTCGTGAGCTGGCTAGAGATTACA 214

QY      61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
      |||||
Db      215 AGTGTCATATTCCTAAAGGTGGATGTGGAGAAATTAAAGAGTGTTCCTCAAGATTGGCT 274

QY      81 IlleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
      |||||
Db      275 ATTGAGGCTATGGCCACTTTTGTGTTTGAAGAGGAGAAAGCTTCTGGCAAAAGTGTG 334

QY      101 GValAlaLysLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
      |||||
Db      335 GGAGCGAAGAGGATGAGCTGCACAGAAATACAGAAACATGTGGCTTCAGCTAGTCT 394

RESULT 3
BG508580      504 bp      mRNA      linear      EST 28-NOV-2001
LOCUS      sac74f08.v1 Gm-cl072 Glycine max cDNA clone GENOME SYSTEMS CLONE
DEFINITION      ID: Gm-cl072-1048 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ;
      mRNA sequence.
ACCESSION      BG508580
VERSION      BG508580.1 GI:13479237
KEYWORDS      EST.
SOURCE      Glycine max (soybean)
ORGANISM      Glycine max
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
      ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
      Glycine.
REFERENCE      1 (bases 1 to 504)
AUTHORS      Shoemaker,R., Keim,P., Vocklin,L., Erpelting,J., Corvelli,V., Khanna
      ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
      Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
      ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
      ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
      ,R., Waterston,R. and Wilson,R.
      Public Soybean EST Project
      Unpublished
      Contact: Shoemaker R/Public Soybean EST Project
      Public Soybean EST Project
      Washington University School of Medicine
      4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
      Tel: 314 256 1800
      Fax: 314 256 1810
      Email: est@watson.wustl.edu
      This clone is available through: ResGen, Invitrogen Corp. 2130
      South Memorial Parkway Huntsville, AL 35801 For further information
      call: (800)-533-4363 or contact via email: ccu@resgen.com
      High quality sequence stop: 439.
FEATURES
      source
      1..504
      /organism="Glycine max"
      /mol_type="mRNA"

```

/notes="Vector: pBluescriptII SK+; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from etiolated hypocotyl tissue of 9-10 day old seedlings of the cultivar Williams 82. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dI) primer with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and XhoI. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."

BASE COUNT 160 a 105 c 134 g 173 t
ORIGIN

Alignment Scores:
Pred. No.: 8.87e-72 Length: 572
Score: 615.00 Matches: 120
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-786-715-6 (-120) x BQ630469 (1-572)

QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20
DB 52 ATGGCTGGCTCATCGAGAGGGAACAGTATTAGTGGCCACCGTTGAGGAATGGAAC 111
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40
DB 112 GATCAACTCCAGAGGGAACGAAATCCAGAACTCATTGTTGTGGATTTCCTGCTCT 171
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
DB 172 TGGTGTGACCCATGCGGTTTCATTGACACCTTCCTTGGCTGGCTGAGGATTCACA 231
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
DB 232 AGTGTCATATTCCTAAAGTGGATGTGGACGAATTAAGAGTGTTCCTCAAGATTGGCT 291
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
DB 292 ATTGAGGCTATGCCACATTTTGTCTTTGTGAAGAGGGAACGCTTCGGACAAAGTGTG 351
QY 101 GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
DB 352 GGAGCAAGAGAGGATGAGCTGCACGACGAGAAATACAGAAACATGTGCTTCAGCTAGT 411

RESULT 5
BQ630383
LOCUS
DEFINITION
saq06f10.y1 Gm-c1045 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1045-4100 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; mRNA
sequence.
ACCESSION
BQ630383

Db 217 AGTGTCATATTCCTAAAGTGGATGTGGACGAATTAAGAGTGTTCCTCAAGATTGGCT 276
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
DB 277 ATTGAGGCTATGCCACATTTTGTCTTTGTGAAGAGGGAACGCTTCGGACAAAGTGTG 336
QY 101 GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
DB 337 GGAGCAAGAGAGGATGAGCTGCACGAGAAATACAGAAACATGTGCTTCAGCTAGTCT 396

RESULT 4
BQ630469
LOCUS
DEFINITION
saq07f11.y1 Gm-c1045 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1045-4125 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; mRNA
sequence.
ACCESSION
BQ630469
VERSION
BQ630469.1 GI:21678108
KEYWORDS
EST.
SOURCE
Glycine max (soybean)
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
1 (bases 1 to 572)
Shoemaker,R., Keim,P., Vodkin,L., Erpelnding,J., Coryell,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
J., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project

TITLE
Unpublished
JOURNAL
Public Soybean EST Project
COMMENT
Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@atson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccudresgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 433.
Location/Qualifiers
1..572
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1045-4125"
/tissue_type="Hypocotyl, 9-10 day old etiolated seedlings"
/lab_host="DH10B"
/clone_lib="Gm-c1045"

FEATURES
source

VERSION BQ630383.1 GI:21678032
 EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE 1 (bases 1 to 582)
 AUTHORS Shoemaker R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
 R., Waterston, R. and Wilson, R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished
 COMMENT Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 296 1800
 Fax: 314 296 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: ccu@resgen.com web site:
 www.resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 430.

FEATURES
 source
 1..582
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl045-4100"
 /tissue_type="Hypocotyl, 9-10 day old etiolated seedlings"
 /lab_host="DH10B"
 /clone_lib="Gm-cl045"
 /note="Vector: pBluescriptII SK+; Site 1: EcoRI; Site 2:
 XhoI; This cDNA library was constructed from mRNA isolated
 from etiolated hypocotyl tissue of 9-10 day old seedlings
 of the cultivar Williams 82. Complementary DNA was
 synthesized from mRNA using a primer consisting of a
 poly(dT) primer with a XhoI restriction site. EcoRI
 adapters were ligated to the blunt-ended cDNA fragments
 followed by digestion with EcoRI and XhoI. The cDNA
 fragments were directionally cloned into the EcoRI-XhoI
 restriction site of the pBluescript vector. The ligated
 cDNA fragments were transformed into DH10B host cells
 (Gibco BRL). This library was constructed by Dr. Randy
 Shoemaker."

BASE COUNT 166 a 105 c 135 g 176 t
 ORIGIN
 Alignment Scores: 9.09e-72 Length: 582
 Pred. No.:

Score: 615.00 Matches: 120
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x BQ630383 (1-582)

Qy 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAla 20
 |||
 Db 52 ATGGCTGGCTCATCGGAGAGGGGCAAGTCATTAGCTGGCCACCGCTGAGSAGTGGAAC 111
 |||
 Qy 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40
 |||
 Db 112 GATCAACTCCAGAAAGGGCAAGGATCCAGAAACTCATTGTGTGATTTTACTGCTTCT 171
 |||
 Qy 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
 |||
 Db 172 TGGTGGGACCATGGCGTTTCATTGCACCATCTTGGCTGAGCTGGCTAAGAGTTCACA 231
 |||
 Qy 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
 |||
 Db 232 AGTGTCATATTCCTAAAGGTGGATGTGGACGAATTAAGAGTGTTTCTCAAGATTGGGCT 291
 |||
 Qy 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
 |||
 Db 292 ATTGAGGCTATGGCCCACTTTTGTGTGTGTAAGAGGGAACCGCTTCGGCAAAAGTGTG 351
 |||
 Qy 101 GlyAlaLysLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAla 120
 |||
 Db 352 GGAGCAAAAGAGGATGAGCTGCACGAAATACAGAAACATGTGGCTTCAGCTAGTCT 411

RESULT 6
 BQ578741
 LOCUS BQ578741
 DEFINITION sar60h12.yl Gm-cl074 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-cl074-6911 5' similar to FR:Q43636 Q43636 THIOREDOXIN. ; mRNA
 sequence.
 ACCESSION BQ578741
 VERSION BQ578741.1 GI:23063968
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max

REFERENCE 1 (bases 1 to 445)
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
 R., Waterston, R. and Wilson, R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished
 COMMENT Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: cu@resgen.com web site:
 www.resgen.com

Seq primer: -40RP from Gibco
 High quality sequence step: 409.

FEATURES

source

1. 445
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl074-6911"
 /tissue_type="seedlings induced for HR (hypersensitive response)"
 /dev_stage="9-11 day old"
 /lab_host="DH103"
 /clone_lib="Gm-cl074"
 /note="Vector: pBluescript II SK+; Site1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 9-11 day old seedlings that were induced for HR (hypersensitive response) by vacuum infiltrating plant tissue with Pseudomonas syringae pv. glycinea carrying the avrB gene (Genetics 141:1597-1604). Plant tissue (expanded unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36, and 53 hrs after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plant care, inoculations, and library construction were performed by Steve Clough (Lila Vodkin lab, University of Illinois)."

BASE COUNT 130 a 81 c 118 g 116 t
 ORIGIN

Alignment Scores:

Pred. Nc.: 2.18e-71 Length: 445
 Score: 611.00 Matches: 119
 Percent Similarity: 99.17% Conservative: 0
 Best Local Similarity: 99.17% Mismatches: 1
 Query Match: 99.35% Indels: 0
 DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x EU578741 (1-445)

QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 79 ATGGCTGGATCATCGGAGAGGGGCAAGTCATTAGCTGCCACACCGTTGAGAGATGGAAC 138
 QY 21 AspGlnLeuGlnLysGlySerGluSerLysLysLeuIleValValAspPheThrAlaSer 40
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 139 GATCAACTCCAGAGGGCAAGAACATCCAGAACTCATTTGTGTGGATTTCATGCTTCT 198
 QY 41 TrpCysGlyProCysArgPheLeuAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 199 TGGTGTGGACCATGGCGTTCTTTGCACCATCTTGGCTGAGCTGGCTAAGAGTTTACA 258
 QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 259 AATGTCTATTTCCTAAAGGTGGATGTGGACGAATTAAGAGTGTTCCTCAGATTGGGCT 318
 QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 319 ATGAGGCTATGCCCCACTTTTGTGTTGAAAGAGGGAACGCTTCCTGCAAAAGTGTGTG 378
 QY 101 GValAlaLysLysAspGluLeuGlnLysLysLysLysHisValAlaSerAlaSerAla 120
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 379 GGAGCAAAAGAGGATGAGCTCCAGCAAAATACAGAAACATGTGGCTTCAGCTATGCT 438

RESULT 7

CA937588/c

LOCUS

DEFINITION

CA937588

VERSION

KEYWORDS

SOURCE

ORGANISM

CA937588

EST

Glycine max (soybean)

Glycine max

Glycine max

1 (bases 1 to 572)

Shoenaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna

,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk

,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

,R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:

www.resgen.com

Possible reversed clone: similarity on wrong strand

Seq primer: -40RP from Gibco

High quality sequence stop: 435.

FEATURES

source

1. .572

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-c1048-6078"

/tissue_type="whole seedlings of greenhouse grown plants"

/dev_stage="1 week old"

/lab_host="DH10B"

/clone_lib="Gm-c1048"

/note="vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The Clark NIL was constructed and seed was provided
by Dr. J. Specht, University of Nebraska (Shoemaker and
Specht, 1995). The cDNA library was constructed from mRNA
isolated from whole seedlings of 1 week old greenhouse
grown plants. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site and a 3' anchor. EcoRI adapters were
ligated to the blunt-ended cDNA fragments followed by XhoI
digestion. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
DH10B host cells (GibcoBRL). The library was constructed
in cooperation with Dr. Paul Reim's laboratory at Northern
Arizona University."

BASE COUNT 174 a 125 c 102 g 171 t

ORIGIN

Alignment Scores:

Pred. No.: 9.03e-70 Length: 572
Score: 600.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.56% Indels: 0
DB: 14 Gaps: 0

US-09-786-715-6 (1-20) x CA937588 (1-572)

QY 4 Sex:SexGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeu 23
|||||
DB 572 TCATCGAGAGGGGACAGTCATTAGCTGCCACACCGTTGAGATGGAGATCACTC 513
QY 24 GlnLysGlyAsnGluSerLysLysLeuLeValValAspPheThrAlaSerTrpCysGly 43
|||||
DB 512 CAGAAAGGCGACGATCCAGAACTCATTGTTGTGGATTCTTCTCTTGTGGA 453
QY 44 ProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIle 63
|||||
DB 432 CCATGGCGTTTCATGCACCATCTTGGCTGAGCTGGGTGAGAGTCCACAGGTGCTATA 393
QY 64 PheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAla 83

DB 392 TTCCTAAAGGTGGAATGGAGCAATTAAAGAGTGTCTCAAGATTGGCTATTGAGCT 333
QY 84 MetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLys 103
DB 352 ATGCCCATCTTTGTGTTGTGAAAGAGGGAAGCGTCTTGACAAAGTGTGGAGCAAG 273
QY 104 LysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
DB 272 AAGGATGAGCTGCAGCAGAAATACAGAAACATGTGGCTTCAGCTAGTGTCT 222

RESULT 8

BE803058

LOCUS

DEFINITION

sequence.

BE803058

VERSION

BE803058.1

GI:10234170

KEYWORDS

SOURCE

Glycine max (soybean)

ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

REFERENCE

1 (bases 1 to 472)

Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna

A., Bella,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

X., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk

R., Ritter,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished

CONTACT: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63105, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Putative full length read

vector to vector length is 622 This clone is available through:

ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL

35801 For further information call: (800)-533-4363 or contact via

email: ccu@resgen.com

High quality sequence stop: 441.

location/Qualifiers

1. .472

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1051-1569"

/tissue_type="floral meristematic mRNA"

/lab_host="DH10B"

/clone_lib="Gm-c1051"

/note="vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from floral meristematic mRNA provided by Dr. Halina Knop of Clemson University. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dI) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 132 a 88 c 116 g 134 t 2 others
ORIGIN

Alignment Scores:

Pred. No.: 1-32e-67 Length: 472
Score: 583.00 Matches: 118
Percent Similarity: 99.16% Conservative: 0
Best Local Similarity: 99.16% Mismatches: 1
Query Match: 94.80% Indels: 1
DB: 10 Gaps: 0

US-09-786-715-6 (1-120) x BE803058 (1-472)

QY 2 AlaGlySerGluGluGlnValIleSerCysHisThrValGluGluTrpAsnAsp 21
|||||
DB 15 GCTGGCTCATCGAAGAGGAGAGTCAATTAGCTGCCACACCGTTGAAGATGGAACGAT 74
QY 22 GlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSerTrp 41
|||||
DB 75 CAATCCAGAGGGCAAGAAATCCAGAACTCATTTGTGTGGATTTTACTGCTTCTTGG 134
QY 42 CysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSer 61
|||||
DB 135 TGTGGACCATGCCGTTTCATTCACCATTTCTGGCTGAGCTGGCTAGAGTTCACAGT 194
QY 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIle 81
|||||
DB 195 GTCATATTCCTAAAGGT-GATGTGGACGAATTAAGAGTGTTCACAGATTGGCTATT 253
QY 62 GluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGly 101
|||||
DB 254 GAGGCTATGCCCACTTTTGTGTTTGTGAAGAGGGAAGCGCTTCGGACAAAGTGGTGGGA 313
QY 102 AlalysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
|||||
DB 314 GCAAGAGAGGATGAGCTGCGAGCAGAGAAATACAGAAACATGTGGCTTCAGCTAGCT 370

RESULT 9

BQ786610
LOCUS BQ786610 522 bp mRNA linear EST 26-JUL-2002
DEFINITION sac70F04.y1 Gm-cl076 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl076-5240 5' similar to FR:Q43636 Q43636 THIOREDOXIN. ; mRNA sequence.
ACCESSION BQ786610
VERSION BQ786610.1 GI:21995082

KEYWORDS

SOURCE Glycine max (soybean)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE

1 (bases 1 to 522)

AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE

Public Soybean EST Project

JOURNAL

Unpublished

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: cou@resgen.com web site: www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 422.

FEATURES

source

1..522
Location/Qualifiers
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl076-5240"
/tissue_type="wounded cotyledons"
/dev_stage="11 day old seedlings"
/lab_host="DH10B"
/clone_lib="Gm-cl076"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 11 day old seedlings treated with that were treated with 2 ugs/ml of a crude glucan elicitor preparation isolated from the mycelial walls of Phytophthora sojae. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plant material was provided by Michael G. Hahn (Complex Carbohydrate Research Center, University of Georgia) and the library was constructed by Anu Khanna (Lila Vodkin Lab, University of Illinois)."

BASE COUNT 142 a 97 c 120 g 163 t
ORIGIN

Alignment Scores:

Pred. No.: 2.05e-67 Length: 522
Score: 582.00 Matches: 113
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.63% Indels: 0
DB: 13 Gaps: 0

US-09-786-715-6 (1-120) x BC786610 (1-522)

Qy 8 GGLnValIleSerCysHisThrValGluGluTrpAsnAspGlnLeuGlnLysGlyAsn 27
Dy 1 GGACAGTCATTAGCTGCCACACCGTTGAAGAAATGGAACGATCACTCAAGAGGCAAC 60
Qy 28 GUSerLysLysLeulleValValAspPheThrAlaSerTrpCysGlyProCysArgPhe 47
Dy 61 GAATCCAGAAACTCATTTGTTGGATTACTGCTTCTTGGTGGACCATGCCGTTTC 120
Qy 48 IleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerValIlePheLeuLysVal 67
Dy 121 ATGACACCATCTTGGCTGAGCTGGCTAAGAGTTCACAGTTCACAGTGTCTATTCCTAAGGTG 180
Qy 68 AspValAspGluLeuLysSerValSerGlnAspTrpAlaIleGluAlaMetProThrPhe 87
Dy 181 GATGGAGCAATTAAGAGTGTCTCAAGATTGGGCTATTGAGGCTATGCCACTTTT 240
Qy 88 ValPheValLysGluGlyThrLeuLeuAspLysValValGlyAlaLysLysAspGluLeu 107
Dy 241 GTCTTTGTAAAGGGAACGCTTCTGGACAAATGTGGTGGAGCAAGAAAGGATGAGCTG 300
Qy 108 GlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
Dy 301 CACAGAAATACAGAAACATGTGGCTTCAGCTAGTGGT 339

RESULT 10

AI988470
LOCUS AI988470 452 bp mRNA linear EST 30-NOV-2001
DEFINITION sd02f07.y1 Gm-cl020 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl020-542 5' similar to TR:Q43636 Q43636 THIOREDOXIN.; mRNA
sequence.

ACCESSION

AI988470

VERSION AI988470.1 GI:5820264

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1 (bases 1 to 452)

AUTHORS

Shoenaker, R., Keim, P., Vockin, L., Erpelting, J., Corvella, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk

R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.

TITLE
JOURNAL
COMMENT

Unpublished
Contact: Shoenaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel.: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 413.

FEATURES

source

1..452
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl020-542"
/tissue_type="root nodules of greenhouse grown plants"
/lab_host="XLI0-Gold"
/clone_lib="Gm-cl020"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from nodules on the roots of 2.5 month-old Glycine max
'Williams' plants that were greenhouse grown. The cDNA
library was prepared using the Stratagene pBluescript II
SK(+) library construction kit. First strand synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An 'anchor'
nucleotide (v=A, C, or G) was added to the 3' end of the
primer [GAGAGAGAGAGAGAGAGACTAGTTCAG(T)18V] to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in with
cloned Pfu DNA polymerase and size- fractionated with a
400 bp cutoff, using a SizeSep 400 Spun column from
Pharmacia. The column eluent was ligated to EcoRI adaptors
and phosphorylated. The XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size- fractionated with a 500 bp cutoff,
using GibcoBRL Life Technologies' cDNA Size Predigested
vector (pBluescript II SK+) that has been digested with
EcoRI and XhoI, and phosphorylated by Stratagene. Both the
white and blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=56) and
sequence (n=16). This library was constructed by Dr. Paul
Keim and Dr. Virgilia Corvella."

BASE COUNT 136 a 80 c 113 g 123 t

ORIGIN

Alignment Scores:

Pred. No.: 2.3e-67 Length: 452

Score: 581.00 Matches: 112

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35901 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 416.

FEATURES

1..466
Location/Qualifiers
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl031-435"
/tissue_type="Williams seedlings, minus the cotyledons"
/lab_host="DH10B"
/clone_lib="Gm-cl031"

/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from whole 'Williams' seedlings, minus the cotyledons, which were propagated on paper towels with distilled water for 5 days, incubated at 40 degrees C for 1 hour. The cotyledons were removed and the remaining tissue was flash-frozen in liquid nitrogen. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGACTAGTCTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (400U/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GibcoBRL Life Technologies' cDNA size fractionation column. The column eluent was then precipitated, redissolved, and ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 100% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=18 and 5, respectively). This library was constructed by Dr. Paul Keim and Dr. Virginia Corvelli."

EASE COUNT 142 a 82 c 116 g 126 t
ORIGIN
Alignment Scores: 2.39e-67 Length: 466
Pred. No.:

Percent Similarity: 97.50% Conservative: 5
Best Local Similarity: 93.33% Mismatches: 3
Query Match: 94.47% Indels: 0
DB: 9 Gaps: 0

US-09-786-715-6 (1-120) * AT988470 (1-452)

QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20
D6 55 ATGGCAGGCTCTCGAGAGGAGGACAGATAGTGTGTACACCGTTGACGATGGAC 114
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeulleValValAspPheThrAlaSer 40
D6 115 GATCACTCCAGAGGCAACCAATCCAGAACTCAITTTGTGTGATTTTACTGCTCT 174
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
D6 175 TGGTGTGGACCATCGCGTTTCATTGACCAATTCCTGGCTGAGCTGGCTAGAAAGTTCACA 234
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
D6 235 AGTGCTGATTCCTAAAGGTGCATGTGGAGAAATTAAGAGTGTTCCTCAAGACTGGCT 294
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
D6 295 ATTGAGGCTATGCCCACTTTTGTGTTTGAAGAGGAAACACTTCAAGCAAGTGGTG 354
QY 101 GlyAlaLysLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
D6 335 GGAGCAAGAGAGATGAGTGCAGCAGACATAGAGAAATATGTGGCTTACGTAGTCT 414

RESULT 11
AW569018 466 bp mRNA linear EST 03-DEC-2001
LOCUS sl74e02.y1 Gm-cl031 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl031-435 5' similar to TR:Q43636 Q43636 THIOREDOXIN. ; mRNA
sequence.

ACCESSION AW569018
VERSION AW569018.1 GI:7233673
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1 (bases 1 to 466)
AUTHORS Shoemaker,R., Keim,P., Vodka,L., Erpelting,J., Corvelli,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.

TITLE Public Soybean EST Project
JOURNAL Unpublished
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project

Score: 581.00 Matches: 112
Percent Similarity: 97.50% Conservatives: 5
Best Local Similarity: 93.33% Mismatches: 3
Query Match: 94.47% Indels: 0
DB: 9 Gaps: 0

US-09-786-715-6 (1-120) x AW569018 (1-466)

QY 1 MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 20
D6 69 ATGGCAGGCTCTCGGAGGAGGCAAGTCTAGCTGTCAACCGTTGACGCAATGGAC 128
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40
D6 129 GATCAACTCCAGAGGCAACCAATCCAGAACTCATTTGTTGGATTTTACTGCTCT 188
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
D6 189 TGTGTGGACCATGCCGTTTCATTGCACCATTTCTTGGCTGAGCTGGCTAAGAGTCA 248
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
D6 249 AGTGTGCTATCTTAAGCTGATGTGAGCAATTAAGAGTGTCTCAAGACTGGCT 308
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal 100
D6 309 ATTGAGGCTATGCCACTTTGTGTTTGAAGAGGCAACACTTCTAAGCAAAAGTGGTG 368
QY 101 GlyAlaLysLysAspGluLeuGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
D6 369 GGAGCAAGAGGATGAGCTGCAGCAACAATAGAGAAATATGTGGCTTAGCTAGTGT 428

RESULT 12
BG838879/c 613 bp mRNA linear EST 25-MAY-2001
LOCUS BG838879
DEFINITION Glycine clandestina cDNA clone Gc01_06f08, mRNA sequence.
ACCESSION BG838879
VERSION BG838879.1 GI:14205201
KEYWORDS EST.
SOURCE Glycine clandestina
ORGANISM Glycine clandestina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 613)
AUTHORS Singh, J.A., Savoie, A.J., Chapados, J., Couroux, P., De Moors, A.,
Harris, L.J., Hattori, J., Ouellet, F., Robert, L.S., Spratt, D. and
Tinker, N.A.
TITLE Differentially Expressed Sequence Tags from Cold-Stressed Glycine
clandestina Seedlings
JOURNAL Unpublished
COMMENT Contact: Singh, J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA

OC6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@agr.ca.
Location/Qualifiers
1. .613
/organism="Glycine clandestina"
/mol_type="mRNA"
/cultivar="1035"
/db_xref="taxon:45687"
/clone="Gc01_06f08"
/tissue_type="Leaves, stem"
/clone_lib="Gc02_AAFRC_EOORC_cold_stressed_Glycine_clandestina SSH"
/notes="Vector: pGEM-T easy; Site 1: EcoRI; Plants incubated at 2 C under 12 hours light/day and harvested after 3 days of cold treatment. SSH (Subtraction Suppressive Hybridization) library made using Clontech PCR-Select cDNA subtraction kit. PCR amplicons were ligated into the pGEM-Teasy vector (Promega)."
BASE COUNT 191 a 145 c 108 g 164 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: 1.18e-66 Length: 6.3
Score: 577.00 Matches: 1.3
Percent Similarity: 95.80% Conservatives: 1
Best Local Similarity: 94.96% Mismatches: 5
Query Match: 93.82% Indels: 0
DB: 12 Gaps: 0

US-09-786-715-6 (1-120) x BG838879 (1-613)

QY 2 AlaGlySerSerGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn 21
D6 553 GCTGGCTCATCGGAGAGGCAAGTCTAGCTGCCACCGTTGAGCATGGAACGAT 494
QY 22 GlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSerTrp 41
D6 493 CAACCTCAAAAGGGCAACGAATCCAGAACTCATTTGTTGGATTTTACTGCTCTGG 434
QY 42 CysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSer 61
D6 433 TGTGGACCATGCCGTTTTCATTGCAACCATTTCTTGGCTGAGCTGGCTAAGAGTCTAGT 374
QY 62 ValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAlaIle 81
D6 373 GTCATATCTTAAGGKGATGTGGACGAATTAAGAGTGTCTCAAGATTGGGCTATT 314
QY 82 GluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValValGly 101
D6 313 GAGGCTATGCCACTTTTGTGTTTGTGAAGAGGCAACGCTCTGGCAAGTGTGGGA 254
QY 102 AlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
D6 253 GCAAGAGGATGAGCTGCAGCAACAATAGAGAAATATGTGGCTTAGCTAGTGTCT 197

RESULT 13
 CD401101/c
 LOCUS 627 bp mRNA linear EST 07-JUN-2003
 DEFINITION Gm_ck2312 Soybean induced by Salicylic Acid Glycine max cDNA 3', mRNA sequence.
 ACCESSION CD401101
 VERSION CD401101.1 GI:31459073
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurossids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 REFERENCE 1 (bases 1 to 627)
 AUTHORS Tian,A.-G., Wang,J., Cui,P., Han,X.-J., Xu,H., Cong,L.-J., Huang,X.-G., Wang,X.-L., Jiao,Y.-Z., Wang,B.-J., Wang,Y.-J., Zhang,J.-S., Chen,S.-Y. and Yu,J.
 TITLE Soybean Expressed Sequence Tags Sequencing
 JOURNAL Unpublished
 COMMENT Contact: Chen S-Y
 Plant Biotechnology Laboratory
 Institute of Genetics and Developmental Biology, CAS, China
 Datun road, Beijing 100101, China
 Tel: 86-10-64886859
 Fax: 86-10-64873428
 Email: sychen@genetics.ac.cn
 Email: sychen@genetics.ac.cn
 Seq primer: T7 primer.
 Location/Qualifiers
 1..627
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Kefeng 1"
 /db_xref="taxon:3847"
 /tissue_type="Seedlings"
 /dev_stage="two-week seedlings"
 /lab_host="XLI-Blue MRF" strain"
 /clone_lib="Soybean induced by Salicylic Acid"
 /note="Vector: pBluescript SK+; Site 1: EcoR I; Site 2: Xho I; The cDNA library was constructed by He, C-Y from mRNA isolated from two-week seedlings (cultivar Kefeng 1) treated by spraying 2.0mM salicylic acid for 24, 36, 48 and 72 h. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XLI-Blue MRF host cells (Stratagene)."
 BASE COUNT 178 a 135 c 111 g 203 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.62e-65 Length: 627
 Score: 566.00 Matches: 118

Percent Similarity: 80.27% Conservative: 0
 Best Local Similarity: 80.27% Mismatches: 0
 Query Match: 92.03% Indels: 29
 DB: 14 Gaps: 1
 US-09-786-715-6 (1-120) x CD401101 (1-627)
 Qy 3 GlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsnAspGln 22
 D6 621 GGTCTCATCGGAGAGGGGACAAAGTCATTAGCTGCCACACCGTTGAAGAATGGAACGATCAA 562
 Qy 23 LeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSerTrpCys 42
 D6 561 CTCCTAAGAGGGCAACGAATCCAAAGAACTCATTTGTTGGATTTCATGCTTCTTGTTGT 502
 Qy 43 GlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThrSerVal 62
 D6 501 GGACATGCGGTTTCATTGCGACCACTCTGTGGCTAGGCTGCTAAGAAGTTCAGAGTGC 442
 Qy 63 IlePheLeuLysValAspValAspGluLeuLys-----73
 D6 441 ATATTCCTAAAGCTGGATGTGGACGAATTAAAGGTAAAAAAAACCCCTACCAAAAAAGCA 382
 Qy 74 -----Se 74
 D6 361 GAGAACTAAGAAAGAAATTCATTAGAAAGTAATCAATAAGATAATTTGTGTGATTCAGAG 322
 Qy 74 rValSerGlnAspTrpAlaIleGluAlaMetProThrPheValPheValLysGluGlyTh 94
 D6 321 TGTTCTCAAGATTGGCTATTGGGCTATGCCACCTTTTGTTGTGAAGAGGGGAC 262
 Qy 94 rLeuLeuAspLysValValGlyAlaLysLysAspGluLeuGlnLysLysIleGlnLysHi 114
 D6 261 GCTTCTGGCAAAAGTGTGGGAGCAAGAGGATGAGCTGCACAGAAATACAGAAACA 202
 Qy 114 sValAlaSerAlaSerAla 120
 D6 201 TGTGCTTCAGCTAGTGTCT 183
 RESULT 14
 CD038084
 LOCUS 687 bp mRNA linear EST 07-MAY-2003
 DEFINITION UPP1002_C04 USDA-Tifton Peanut Immature Pod cDNA Library (UTP)
 ACCESSION CD038084
 VERSION CD038084.1 GI:30419922
 KEYWORDS EST.
 SOURCE Arachis hypogaea (peanut)
 ORGANISM Arachis hypogaea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurossids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.
 REFERENCE 1 (bases 1 to 687)
 AUTHORS Luc,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G. and Lynch,R.E.
 TITLE Generation and Analyses of ESTs for Arachis hypogaea

JOURNAL COMMENT

Unpublished
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 228-387-2334
Email: bguo@tifton.usda.gov
Seq primer: T3.

FEATURES

source
1. 687
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="JTPP1002 C04"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="X11-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent ultrapur(GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

BASE COUNT
ORIGIN

Alignment Scores:

Pred. No.: 8.92e-64 Length: 687
Score: 556.00 Matches: 103
Percent Similarity: 96.67% Conservative: 13
Best Local Similarity: 85.93% Mismatches: 4
Query Match: 90.41% Indels: 0
DB: 14 Gaps: 0

US-09-786-715-6 (1-120) x CD038084 (1-687)

QY 1 MetAlaGlySerSerGluGluGlnValIleSerCysHisThrValGluGluTrpAsn 20
|||||
Db 69 ATGGCTGGTTCATCAGAGAGGGACAGATTATCGGTTCCACTCGGTGATGATGGACC 128
|||||
QY 21 AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer 40

Db 129 CAACAGCTCGAGAGGGAAATGAATCTAAGAGCTGATTTGTTGGATTCTACTGCTCA 168
|||||
QY 41 TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysLysPheThr 60
|||||
Db 189 TGGTGGGACCATGGCGCTTCATTGCACCTTTTCTGGCTGAGTTGGCTAAGAAGTTTACA 248
|||||
QY 61 SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla 80
:::|||||
Db 249 AATGTCATATTTTGAAGTGGATGTGATGAAGTCACTACATCTGTTGCTCAAGACTGGGCT 308
|||||
QY 81 IleGluAlaMetProThrPheValPheValLysGluGlyThrIleLeuAspLysValVal 100
:::|||||
Db 309 GTGGAGCAATGCCACCTTTGTGTTTGTGAAGAGGAGCAACATTATTGCAAAAGTGTG 368
|||||
QY 101 GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla 120
|||||
Db 369 GGAGCAAGAAAGGAAGACTGCAGCAGAAAGCTTGAGAAACAATGGCTACAGCCAGT 428
|||||

RESULT 15

CA901886 570 bp mRNA linear EST 27-DEC-2002
CA901886
LOCUS
DEFINITION
PSCC14170 Scarlet Runner Bean Suspensor Region Triplex2 Phaseolus
coccineus cDNA 5' similar to Thiofodoxin H-type, mRNA sequence.
ACCESSION
CA901886
VERSION
CA901886.1 GI:27388878
KEYWORDS
EST.
SOURCE
Phaseolus coccineus
ORGANISM
Phaseolus coccineus

REFERENCE
1 (bases 1 to 570)
Bui,A.Q., Le,B.H., Weterings,K., Bi,Y.-P., Choi,J.-S., McElroy,K.E.,
Choi,P.S., Harada,J.J., Fischer,R.L. and Goldberg,R.B.
Gene Activity in Different Regions of a Post-Fertilization Plant
Embryo by EST Analysis
Unpublished
Contact: Goldberg, R.B.
Department of Molecular, Cell, & Developmental Biology
University of California, Los Angeles
621 Charles E. Young Drive South, Los Angeles, CA 90095-1606, USA
Tel: 310 825 3270
Fax: 310 825 8201
Email: bobg@ucla.edu
Seq primer: 5' Triplex
POLYA-No.

FEATURES

1. 570
/organism="Phaseolus coccineus"
/mol_type="mRNA"
/cultivar="Hammond's Dwarf Scarlet"
/db_xref="taxon:3886"
/dev_stage="6-days post-pollination"
/clone_lib="Scarlet Runner Bean Suspensor Region Triplex2"
/note="Organ: Suspensor Region of Globular-Stage Embryos;

Vector: Triplex2; Site 1: SfiI; Site 2: SfiI; Suspensor regions were micro-dissected from globular-stage embryos six days after pollination from greenhouse-grown plants [Weterings et al., Plant Cell 13, 2409-2425 (2001)]. Double-stranded cDNA was synthesized from suspensor mRNA using the SMART cDNA library construction kit according to the manufacturer (Clontech). The suspensor cDNA fragments were directionally ligated into the SfiI restriction site of the lambda Triplex2 vector (Clontech), and the recombinant cDNAs were transformed into E. coli XL1-Blue cells (Clontech). Suspensor cDNA plasmids used for directional sequencing were obtained by in vivo excision from the lambda Triplex2 recombinants in E. coli BM25.8 cells (Clontech)."

BASE COUNT 154 a 92 c 134 g 190 t
ORIGIN

Alignment Scores:
Pred. No.: 8.31e-61 Length: 570
Score: 533.00 Matches: 100
Percent Similarity: 95.00% Conservative: 14
Best Local Similarity: 83.33% Mismatches: 6
Query Match: 86.67% Indels: 0
DB: 14 Gaps: 0

US-09-786-715-6 (1-120) x CA901896 (1-570)

QY	1	MetAlaGlySerSerGluGluGlyGlnValIleSerCysHisThrValGluGluTrpAsn	20
Db	82	ATGCGCGCGCATCAGAGAGGACAGTCAATGGCTGCCACACCGTTGAGGCATGGACA	141
QY	21	AspGlnLeuGlnLysGlyAsnGluSerLysLysLeuIleValValAspPheThrAlaSer	40
Db	142	GAACACTCGAAAGGCAATATCCAGAAACTAATGTGTGATTTACTGCTTCT	201
QY	41	TrpCysGlyProCysArgPheIleAlaProPheLeuAlaGluLeuAlaLysPheThr	60
Db	202	TGGTGGGACCAATGCCGTTTCATTTCTCCATTTGGCTGAGCTGGCTAAGAGTTTACA	261
QY	61	SerValIlePheLeuLysValAspValAspGluLeuLysSerValSerGlnAspTrpAla	80
Db	262	AATGCTGTAATTTCTGAAGTGGATGTAGACGATTAAAGAGTGTTCGTAAGATTTTGT	321
QY	81	IleGluAlaMetProThrPheValPheValLysGluGlyThrLeuLeuAspLysValVal	100
Db	322	GTGAGCAATGCCCACTTTTGTGTTTGAAAGAGGAACTCTCTGGCAAAAGTGTG	381
QY	101	GlyAlaLysLysAspGluLeuGlnGlnLysIleGlnLysHisValAlaSerAlaSerAla	120
Db	382	GGACAAAGAGGAGAGATTCGACGACAAATAGAGAAACATGTGCTGCACCTAGTGT	441

Search completed: December 5, 2003, 13:00:41
Job time : 2064 secs